

Abstract

The present invention provides a method and apparatus for increasing the performance of world-wide-web traffic over the Internet. A distributed network of specialized nodes of two types is dispersed around the Internet. A web client's requests are directed to a node of the first type chosen to be close to the client, and the client communicates with this node using a standard protocol such as HTTP. This first node receives the request, and communicates the request to a node of the second type chosen to be close to the request's ultimate destination (e.g., a web server capable of generating a response to the request.) The first node communicates the request to the second node using a different, specialized, protocol that has been designed for improved performance and specifically to reduce traffic volume and to reduce latency. The second node receives communication from the first node using this specialized protocol, converts it back to a standard protocol such as HTTP, and forwards the request to the destination computer or server. Responses from the destination to the client take the corresponding reverse route, and also are carried over a specialized protocol between the two nodes. In addition, these nodes can employ other techniques such as web caches that avoid or improve some communication steps. Thus, specialized, proprietary, or complex protocols and techniques can be quickly deployed to enhance web performance without requiring significant changes to the clients or servers.